

COMPUTABLE RANDOMNESS AND MONOTONICITY

ALEX GALICKI

ABSTRACT. We show that $z \in \mathbb{R}^n$ is computably random if and only if every computable monotone function on \mathbb{R}^n is differentiable at z .

1. INTRODUCTION

Our main result is concerned with differentiability of monotone functions of several variables. Monotone functions are closely related to Lipschitz functions and they play a prominent role in variational analysis (see [10]) and in the theory of optimal transport (see [12]). It is known that on the unit interval differentiability of computable monotone functions is equivalent to computable randomness.

Theorem 1.0.1 (Theorem 4.1 in [3]). *A real z is computably random \iff every computable nondecreasing function $f : [0, 1] \rightarrow \mathbb{R}$ is differentiable at z .*

We will prove the following generalization of the above result.

Theorem 1.0.2. *Let $n \geq 1$. $z \in \mathbb{R}^n$ is computably random \iff every computable monotone function $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is differentiable at z .*

The proof has two distinct parts. The \Rightarrow implication is proven in Section 2 using an effective form of the Rademacher Theorem and geometric properties of monotone functions. The converse implication is proven in Section 3 and uses results from optimal transport.

2. DIFFERENTIABILITY OF COMPUTABLE MONOTONE FUNCTIONS FROM \mathbb{R}^n TO \mathbb{R}^n

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a function. We say f is *monotone* if

$$\langle f(x) - f(y), x - y \rangle \geq 0 \text{ for all } x, y \in \mathbb{R}^n.$$

As we will see later in this section, monotone functions are very closely related to Lipschitz functions.

In non-effective setting, a.e. differentiability of monotone function from \mathbb{R}^n to \mathbb{R}^n has been proven by Mignot [8], who used Rademacher's Theorem and a fact about monotone functions discovered by Minty [9].

In this section we will show that computable randomness implies differentiability of computable monotone real functions of several variables and our proof follows the same path - using the effective form of Rademacher's Theorem proven in the previous section and the following correspondence observed by Minty.

2.1. Minty parameterization and overview of the proof. Minty showed that the so called Cayley transformation

$$\Phi : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n \times \mathbb{R}^n \text{ defined by } \Phi(x, y) = \frac{1}{\sqrt{2}}(y + x, y - x)$$

transforms the graph of a monotone function into a graph of a 1-Lipschitz function. Note that when $n = 1$ this is a clockwise rotation of $\pi/4$. We will rely on the following consequence of the above fact.

Proposition 2.1.1 (cf. Proposition 1.2 in [1]). *Let $u : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be monotone. Then $(u + I)$ and $(u + I)^{-1}$ are monotone and $(u + I)^{-1}$ is 1-Lipschitz.*

Proposition 2.1.2 (cf. Theorem 12.65 in [10]). *Let $u : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a continuous monotone function. Let $z \in \mathbb{R}^n$ and define $f = (u + I)^{-1}$ and $\hat{z} = u(z) + z$. The following two are equivalent:*

- (1) *u is differentiable at z , and*
- (2) *f is differentiable at \hat{z} and $f'(\hat{z})$ is invertible.*

A good exposition of classical results related to this area can be found in [1] and [10].

Now we are ready to explain our proof.

Overview of the proof

Let $u : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a monotone computable function and let $z \in [0, 1]^n$ be computably random. Then $g = u + I$ is monotone and computable and $f = g^{-1}$ is 1-Lipschitz and computable. If we can show that $g(z)$ is computably random, then f is differentiable at $g(z)$. By Proposition 2.1.2, if the derivative of f at $g(z)$ is invertible, then g is differentiable at z .

From the above description, it is clear that we require the following two ingredients to complete the proof:

- (preservation property) we need to show that $g(z)$ is computably random when z is, and that
- (singularity property) computable randomness of $g(z)$ implies that $f'(g(z))$ is invertible.

In the following two subsections we will prove both of the above.

2.2. Another preservation property. To prove the preservation property mentioned in the previous subsection, we require some terminology and notation from [?].

Firstly, we need to extend the notion of computable randomness to \mathbb{R}^n : we say $z \in \mathbb{R}^n$ is computably random if its binary expansion (or, equivalently, its fractional part) is computably random. When $z \in [0, 1]^n$, this characterisation is equivalent to the Definition ???. Otherwise, when $z \notin [0, 1]^n$, this characterisation is equivalent to computable randomness on some computable translation of the unit cube equipped with the usual Lebesgue measure.

Notation 2.2.1. For every $n \geq 1$, let \mathcal{A}_n be some fixed a.e. decidable cell decomposition of $[0, 1]^n$. For the sake of simplifying the notation, in the rest of this section, for all $n \geq 1$ and all $\sigma \in 2^{<\omega}$, we denote the cell $[\sigma]_{\mathcal{A}_n}$ by $[\sigma]$.

Definition 2.2.2. A *Martin-Löf test* is a uniformly computable sequence $(U_i)_{i \in \mathbb{N}}$ of Σ_1^0 subsets of $[0, 1]^n$ such that $\lambda(U_i) \leq 2^{-i}$ for all i . We say $(U_i)_{i \in \mathbb{N}}$ *covers* $z \in [0, 1]^n$ if $z \in \bigcap_i U_i$.

We say a Martin-Löf test $(U_i)_{i \in \mathbb{N}}$ is *bounded* if there is a computable measure $\nu : 2^{<\omega} \rightarrow [0, \infty)$ satisfying

$$\lambda(U_i \cap [\sigma]) \leq 2^{-i} \nu(\sigma)$$

for all $i \in \mathbb{N}$ and $\sigma \in 2^{<\omega}$.

We require the following characterisation of computable randomness in the unit cube due to Rute:

Proposition 2.2.3 (cf. Theorem 5.3 in [?]). *Let $z \in [0, 1]^n$. The following two are equivalent:*

- (1) *z is not computably random, and*
- (2) *either z is an unrepresented point, or there is a bounded Martin-Löf test $(U_i)_{i \in \mathbb{N}}$ that covers z .*

Remark 2.2.4. For our considerations it is sufficient to know that if z is an unrepresented point, then it is not weakly random.

We are now in position to state and to prove the required preservation property for computable randomness.

Lemma 2.2.5. *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a computable injective Lipschitz function and suppose $z \in \mathbb{R}^n$ is not computably random. Then $f(z)$ is not computably random either.*

Proof. Without loss of generality we assume $z \in [0, 1]^n$, $f(z) \in [0, 1]^n$ and $\mathbf{Lip}(f) \leq 1$ (otherwise we may consider $\hat{f}(x) = A \cdot f(x) + B$ for some suitable computable A and B).

Firstly, let's assume that z is an unrepresented point. Let $P \subset [0, 1]^n$ be a Π_1^0 null set with $z \in P$. Then $f(z) \in f(P) \cap [0, 1]^n$ and since $f(P) \cap [0, 1]^n$ is also a Π_1^0 null set, $f(z)$ is not weakly random.

Let $(V_i)_{i \in \mathbb{N}}$ be a bounded Martin-Löf test with $z \in \bigcap_i V_i$ and let ν be a computable measure such that $\lambda(V_i \cap [\sigma]) \leq 2^{-i} \nu(\sigma)$ for all i, σ .

Define $U_i = f(V_i) \cap [0, 1]^n$ for all i . Since $\lambda(U_i) \leq \lambda(V_i)$ (see Lemma 3.10.12 in [2]) and f is injective, $(U_i)_{i \in \mathbb{N}}$ is a Martin-Löf test.

Define $\nu_f = \nu \circ f^{-1}$. It is a computable measure and for all i, σ we have

$$\begin{aligned} \lambda(U_i \cap [\sigma]) &= \lambda(f(V_i) \cap [\sigma]) = \\ \lambda(f(V_i \cap f^{-1}([\sigma]))) &\leq 2^{-i} \nu(f^{-1}([\sigma])) = 2^{-i} \nu_f(\sigma). \end{aligned}$$

It follows that $(U_i)_{i \in \mathbb{N}}$ is a bounded Martin-Löf test that covers $f(z)$ and thus $f(z)$ is not computably random. □

2.3. Singularity property. The main result in this subsection, Theorem 2.3.2, can be seen as an effective version of Sard's Theorem for Lipschitz function. Its classical version, proven by Mignot ([8], also see Theorem 9.65 in [10]), states that for a Lipschitz function $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, the set of its critical values is a null-set.

Lemma 2.3.1. *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a Lipschitz function. Suppose $z \in \mathbb{R}^n$ is such that $f'(z)$ is singular. Then for every $\epsilon > 0$, there exists an open neighbourhood of z , O_ϵ such that $\lambda(f(O_\epsilon)) \leq \epsilon \lambda(O_\epsilon)$.*

Proof. Fix $\epsilon > 0$ and let $k = \mathbf{Lip}(f)$.

Define $\epsilon' = \frac{\epsilon}{k^{n-1}2^n(\sqrt{n})^n}$. Since f is differentiable at z , there exists $\delta > 0$ such that

$$|f(x) - f(z) - f'(z)(x - z)| \leq \epsilon'|x - z| \quad (1)$$

for all $x \in \mathbb{R}^n$ with $|x - z| \leq \delta$. There is an open n -cube C with side length equal to $s = \frac{\delta}{\sqrt{n}}$ such that $z \in C$ and 1 holds for all $x \in C$.

Let L be the mapping defined by $L(x) = f(z) + f'(z)(x - z)$. Since $f'(z)$ is singular, L is not onto and its range is contained in some hyperplane H .

As a consequence of 1 we have $|f(x) - L(x)| \leq \epsilon'\delta$ for all $x \in C$. Thus, $f(C) \subseteq L(C) + [-\epsilon'\delta, \epsilon'\delta]^n$. Since L is a k -Lipschitz mapping, the image of C under L lies in the intersection of H with a closed ball with radius $k\delta$ centered at $f(z)$. Then $L(C)$ is contained in a rotated $(n - 1)$ -dimensional cube of side $2k\delta$. This shows that $f(C)$ lies in a rotated box \hat{C} with

$$\lambda(\hat{C}) = (2k\delta)^{n-1}2\epsilon'\delta = 2(2k)^{n-1}\epsilon'(\sqrt{n})^n \left(\frac{\delta}{\sqrt{n}}\right)^n = \epsilon\lambda(C).$$

□

Theorem 2.3.2. *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a computable Lipschitz function and let $z \in \mathbb{R}^n$. If $f(z)$ is computably random, then $f'(z)$ is not singular.*

Proof. Without loss of generality we may assume $f(z) \in [0, 1]^n$ and $[0, 1]^n \subseteq f([0, 1]^n)$. The proof is by contraposition. Suppose $f'(z) = 0$.

Let $\nu = \lambda \circ f^{-1}$ and for every $i \in \mathbb{N}$, define $V_i \subset [0, 1]^n$ as the union of all $[\sigma]$ such that $\lambda(\sigma) \leq 2^{-i}\nu(\sigma)$. Note that $\lambda(V_i) \leq 2^{-i}$ and for every τ ,

$$\lambda(V_i \cap [\tau]) = \sum_{[\eta] \subseteq [\tau] \cap V_i} \lambda(\eta) \leq 2^{-i} \sum_{[\eta] \subseteq [\tau] \cap V_i} \nu(\eta) \leq 2^{-i}\nu(\tau).$$

Thus $(V_i)_{i \in \mathbb{N}}$ is a bounded Martin-Löf test and, by Lemma 2.3.1, it covers $f(z)$.

□

2.4. Main result. We are now ready to formulate and prove the main result of this section.

Theorem 2.4.1. *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be an computable monotone function and let $z \in [0, 1]^n$ be computably random. Then f is differentiable at z .*

Proof. Define $g = (f + I)^{-1}$, then g is a computable Lipschitz function with $\mathbf{Lip}(g) \leq 1$.

Let $y = f(z) + z$ so that $g(y) = z$. By Lemma 2.2.5, y is computably random and hence g is differentiable at y and by Theorem 2.3.2 $g'(y)$ is invertible. Hence, by Proposition 2.1.2, f is differentiable at z .

□

3. MONOTONE TRANSFER MAPS

Suppose $z \in \mathbb{R}^n$ is not computably random and we want to exhibit a computable monotone function $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ that is not differentiable at z . Let us first overview how this problem has been resolved in the case when $n = 1$.

Example 3.0.2 (On the real line). Suppose Z is the binary expansion of z . We start with a martingale M (we may assume it has the saving property) that succeeds on the Z and define a computable measure on the real line by $\mu_M([\sigma]) = M(\sigma) \cdot 2^{-|\sigma|}$. Then the cumulative distribution of μ_M , $f = \mathbf{cdf}_{\mu_M}$, is not differentiable at z .

Before proceeding to generalize this construction in \mathbb{R}^n , we need to review some basic notions from the area known as *optimal transport*.

3.1. Optimal transportation. Let μ, ν be probability measures on \mathbb{R}^n . A probability measure π on $\mathbb{R}^n \times \mathbb{R}^n$ is said to have *marginals* μ and ν when the following holds for all measurable $A, B \subseteq \mathbb{R}^n$:

$$\pi[A \times \mathbb{R}^n] = \mu[A], \text{ and } \pi[\mathbb{R}^n \times B] = \nu[B].$$

Let $\Pi(\mu, \nu)$ denote the set of all probability measures on $\mathbb{R}^n \times \mathbb{R}^n$ whose marginals are μ and ν . Note that this set is always nonempty. For a given *cost function* $c : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ and $\pi \in \Pi(\mu, \nu)$, define the *total transportation cost* $I_c[\pi]$ as

$$I_c[\pi] = \int_{\mathbb{R}^n \times \mathbb{R}^n} c(x, y) d\pi(x, y).$$

The *optimal transportation cost* between μ and ν is the value

$$\mathbb{I}_c(\mu, \nu) = \inf_{\pi \in \Pi(\mu, \nu)} I_c[\pi].$$

Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a map. We say T is a *transport map*, or that T *transports* μ onto ν (in symbols, $\nu = T\#\mu$), if for all measurable A , $\lambda(A) = \mu(T^{-1}(A))$.

Elements of $\Pi(\mu, \nu)$ are called *transference plans*. We are interested in transference plans induced by measurable maps, that is, plans of the form $\pi_T = (I \times T)\#\mu \in \Pi(\mu, \nu)$ where $T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a measurable map. The total transportation cost associated with a transport map T is

$$I_c[T] = I_c[\pi_T] = \int_{\mathbb{R}^n} c(x, T(x)) d\mu(x).$$

A transport map T for which the cost is optimal, that is for which $I_c[\pi_T] = \mathbb{I}_c(\mu, \nu)$, is called an *optimal transport map*. The problem of minimizing $I_c[T]$ over the set of all transfer maps is known as *Monge's optimal transportation problem*.

Let $U \subseteq \mathbb{R}^n$ be compact. For a function $f : U \rightarrow \mathbb{R}$, define its *convex conjugate* f^* by

$$f^*(y) = \sup_{x \in U} [xy - f(x)].$$

Since U is assumed to be compact, f^* is computable when f is.

The following important result lies at the heart of our construction.

Theorem 3.1.1 (Brenier's theorem, cf. Theorem 2.12 in [12]). *Let μ, ν be probability measures on \mathbb{R}^n . Suppose μ is absolutely continuous (with respect to the Lebesgue measure) and the following holds:*

$$\int_{\mathbb{R}^n} \frac{|x|^2}{2} d\mu(x) + \int_{\mathbb{R}^n} \frac{|y|^2}{2} d\nu(y) < \infty.$$

Then there exists a convex function ϕ such that $\nabla\phi\#\mu = \nu$. Moreover, $\nabla\phi$ is the unique (i.e. uniquely determined μ -almost everywhere) gradient of a convex function which pushes μ forward to ν .

Furthermore, if ν absolutely continuous, then, for μ -almost all x and for ν -almost all y ,

$$\nabla\phi^* \circ \nabla\phi(x) = x, \quad \nabla\phi \circ \nabla\phi^*(y) = y,$$

and $\nabla\phi^*$ is the (ν -almost everywhere) unique gradient of a convex function which pushes ν forward to μ , and also the solution of the Monge problem for transporting ν onto μ with a quadratic cost function.

Remark 3.1.2. The above result is known to hold not only for absolutely continuous measures, but more general results are not needed in this paper.

Now we are ready to review the Example 3.0.2 in the context of optimal transport theory.

3.2. The main idea. Let $f = \mathbf{cdf}\mu_M$ be the function from the example. Note that f is a transport map from μ_M to the Lebesgue measure λ (that is, $\lambda = f\#\mu_M$). In fact, by the optimal transportation theorem for a quadratic cost of \mathbb{R} (see Theorem 2.18 in [12]), f is the (unique) optimal transport map from μ_M to λ . Unlike in higher dimensions, on the real line, the form of the optimal transport map is known and in our case (a special case of transporting μ_M onto λ), the function f is the optimal one.

Note that the derivative $D_{\lambda}\mu_M(z)$ of μ_M with respect to the Lebesgue measure does not exist. Intuitively, μ_M oscillates around z and, correspondingly, the transport map is not differentiable at z .

3.3. Wasserstein metrics. Given a Polish metric space (X, d) , the set $\mathcal{P}(X)$ of Borel probability measures over X endowed with the weak topology is a Polish space.

Suppose $(X, d, (\alpha_i)_{i \in \mathbb{N}})$ is a computable metric space where d is bounded. Let $(\delta_i)_{i \in \mathbb{N}}$ be an effective enumeration of those elements of $\mathcal{P}(X)$ which are concentrated on finite subsets of special points and assign rational values to them. Let π be the Prokhorov metric on $\mathcal{P}(X)$, then $(\mathcal{P}(X), \pi, (\delta_i)_{i \in \mathbb{N}})$ is a computable metric space compatible with the weak topology on $\mathcal{P}(X)$. Following [5] and [6], we define computable measures as computable elements of $(\mathcal{P}(X), \pi, (\delta_i)_{i \in \mathbb{N}})$.

For $p \in \mathbb{N}$ with $p \geq 1$, define the cost function c_p by $c_p(x, y) = d(x, y)^p$. For $\mu, \nu \in \mathcal{P}(X)$, define the *Wasserstein metric of order p* by

$$W_p(\mu, \nu) = \mathbb{I}_p(\mu, \nu)^{1/p}$$

where \mathbb{I}_p is the optimal transport cost between μ and ν with respect to c_p . It is known that W_p metrizes the weak topology on $\mathcal{P}(X)$. Furthermore, W_1 is computable and it is *computably equivalent* to π [6]. That is, given a Cauchy name of μ with respect to π , it is possible to compute a Cauchy name of μ with respect to W_1 and vice versa. Since we are mainly concerned with the quadratic cost, we need to prove an analogous result for $p > 1$.

Proposition 3.3.1. *Let $(X, d, (\alpha_i)_{i \in \mathbb{N}})$ be a computable metric space where d is bounded. Let $p > 1$. Then W_p is computably equivalent to W_1 (and hence to π).*

Proof. Firstly, note that $W_p(\delta_i, \delta_j)$ is computable uniformly in i, j, p . This is due to the fact, that computing W_p between discrete measures is a linear programming problem, for which there are algorithms available. (**TODO: some refs**)

It is known (see 7.1.2 in [12]) that the following inequalities hold:

$$W_1 \leq W_p \leq W_1^{1/p} \text{diam}(X)^{1-1/p}. \quad (2)$$

Fix a computable real D with $D \geq \text{diam}(X)$. Let $(\mu_i)_{i \in \mathbb{N}}$ be Cauchy names of μ with respect to π . Let $i, j \in \mathbb{N}$. Using 2 and the triangle inequality we have

$$\begin{aligned} W_p(\mu, \delta_j) &\leq W_p(\delta_j, \mu_i) + W_p(\mu_i, \mu) \leq \\ &D \cdot W_1^{1/p}(\mu, \mu_i) + W_p(\delta_j, \mu_i). \end{aligned}$$

This shows that we can effectively find a Cauchy name with respect to W_p given a Cauchy name with respect to π .

For the other direction, suppose $(\hat{\mu}_i)_{i \in \mathbb{N}}$ is a Cauchy name of μ with respect to W_p . Then

$$\begin{aligned} W_1(\mu, \delta_j) &\leq W_1(\delta_j, \hat{\mu}_i) + W_1(\hat{\mu}_i, \mu) \leq \\ &W_1(\delta_j, \hat{\mu}_i) + W_p(\hat{\mu}_i, \mu). \end{aligned}$$

The required result follows. \square

Corollary 3.3.2. $\mathbb{I}_p(\mu, \nu)$ is computable uniformly in μ, ν and p .

3.4. An effective version of Brenier's theorem.

Theorem 3.4.1. Let μ, ν be absolutely continuous computable probability measures on \mathbb{R}^n with $\text{supp}(\mu) = [0, 1]^n$. There exists a computable convex function $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ such that $\nabla \phi$ is the optimal transport map from μ to ν .

Proof. From Theorem 3.1.1 we know that there is a unique convex function ϕ such that $\phi(0) = 0$ and $\nabla \phi$ is the optimal transformation map from μ to λ . Since it doesn't matter how ϕ is defined outside of $\text{supp}(\mu)$, we may assume ϕ is Lipschitz. Pick some rational $K \in \mathbb{Q}$ so that $K > \mathbf{Lip}(\phi)$ and consider the subspace

$$L_0(K) = \{f \in C[0, 1]^n : \mathbf{Lip} f \leq K \text{ and } f(0) = 0\}.$$

By Arzela-Ascoli theorem, $L_0(K)$ is a compact subspace of $C[0, 1]^n$ (the space of real valued continuous functions endowed with the supremum metric) containing ϕ . Moreover, since the support of μ is equal to $[0, 1]^n$ and $\nabla \phi$ is uniquely determined μ -a.e., ϕ is the only function in $L_0(K)$ for which $\nabla \phi$ is optimal.

Recall, that a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is called *piecewise affine* if there exists a finite set of affine functions $f_i(x) = A_i \cdot x + b_i$, $i = 1, \dots, k$, such that the inclusion $f(x) \in \{f_i(x), \dots, f_k(x)\}$ holds for all x . The functions f_i are called *selection functions*. The set of pairs (A_i, b_i) is called a collection of *matrix-vector pairs* corresponding to f . If f is a piecewise affine function and (A_i, b_i) for $i \leq k$ are the corresponding matrix-vector pairs, there exists a finite number of index sets $M_1, \dots, M_l \subseteq \{1, \dots, k\}$ such that

$$f(x) = \max_{1 \leq i \leq l} \min_{j \in M_i} A_j \cdot x + b_j \text{ for all } x.$$

For every $i \in \mathbb{N}^+$, let \mathcal{D}_i^n denote the set of points in \mathbb{R}^n with all coordinates of the form $k2^{-i}$ for some integer k .

It is known that piecewise affine functions are dense in $L_0(K)$. For $k \in \mathbb{N}^+$, let Γ_k be the (finite) set of piecewise affine functions f such that

- (1) all of its matrix-vector pairs belong to $\mathcal{D}_k^n \times D_k^1$,
- (2) $\mathbf{Lip}(f) \leq K$, and

(3) $f(0) = 0$.

Note that it is possible to effectively enumerate elements of Γ_k uniformly in k . Let $(\gamma_i)_{i \in \mathbb{N}}$ an effective enumeration of $\bigcup_k \Gamma_k$. It is dense in $L_0(K)$.

Claim 3.4.2. $L = (L_0(K), \|\cdot\|_\infty, (\gamma_i)_{i \in \mathbb{N}})$ is an effectively compact computable metric space.

Proof. L is clearly a computable metric space. To show that it is effectively compact, fix $i \in \mathbb{N}^+$. Then $\{B(\gamma; 2^{-i}) \mid \gamma \in \Gamma_i\}$ form a finite open cover of L . \square

For any $f, g : \mathbb{R}^n \rightarrow \mathbb{R}$ let

$$J(f, g) = \int_{\mathbb{R}^n} f(x) d\mu(x) + \int_{\mathbb{R}^n} g(x) d\lambda(x).$$

We know that $J(\phi, \phi^*) = \mathbb{I}_2(\mu, \lambda)$ (see the proof of Theorem 2.12 in [12]).

Define

$$S = \{f \in L_0(K) \mid J(f, f^*) = \mathbb{I}_2(\mu, \lambda)\}.$$

The condition $J(f, f^*) = \mathbb{I}_2(\mu, \lambda)$ is (uniformly) computable in f and μ . Hence S is a Π_1^0 subset. Since ϕ is uniquely defined μ -a.e., S contains only one element - ϕ .

Let us show that effective compactness of $L_0(K)$ guarantees computability of ϕ . The set A_S of basic open balls disjoint from S is recursively enumerable. As we have shown, for a given $j \in \mathbb{N}$, we can find a finite cover of $L_0(K)$ by basic open 2^{-j} -balls. Let us denote such covers I_j . Fix $i \in \mathbb{N}$. Enumerate elements of A_S and elements of those I_j , where $j \geq i$, until all the balls in I_{i+3} that has not been enumerated so far have centers at most 2^{-i-1} from each other. Let γ be one such center. Then the basic open ball $B(\gamma; 2^{-i})$ contains S . Therefore, ϕ is computable. \square

3.5. Application to computable randomness on \mathbb{R}^n . In this subsection we prove the following converse to the Theorem 2.4.1.

Theorem 3.5.1. *Suppose $z \in \mathbb{R}^n$ is not computably random. Then there exists a computable monotone function $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ not differentiable at z .*

We may assume that every coordinate of z is computably random. For suppose z_i is not computably random for some i . There exists a computable monotone function $g : \mathbb{R} \rightarrow \mathbb{R}$ not differentiable at z_i . Then the function $(x_1, \dots, x_n) \mapsto (g(x_1), \dots, g(x_n))$ is a computable monotone function from \mathbb{R}^n to \mathbb{R}^n not differentiable at z .

Since z is not computably random, there exists an absolutely continuous computable probability measure μ on \mathbb{R}^n such that $D_\lambda \mu(z)$ does not exist. This follows from Theorem 5.3(4) [11] and the fact that all coordinates of z are computably random (and hence non-dyadic). Without loss of generality we may assume that μ is supported on $[0, 1]^n$. The following classical result is needed to show that the optimal transport map (from μ onto λ) is not differentiable at z .

Theorem 3.5.2 (Jacobian theorem for monotone maps, cf. Theorem A.2 in [7]). *Let ϕ be a convex function on \mathbb{R}^n and suppose it is twice differentiable at $x \in \mathbb{R}^n$. Then*

$$\lim_{r \rightarrow 0} \frac{\lambda(\partial\phi(B_r(x)))}{\lambda(B_r(x))} = \det D_A^2 \phi(x).$$

Theorem 3.1.1 shows that there exists a unique monotone function $f = \nabla\phi$, such that $f\#\mu = \lambda$, where ϕ is some convex function (obviously, not unique). Since μ is absolutely continuous, by Theorem 2.12 (iv) and Lemma 4.6 from [12], $\lambda(\partial\phi(A)) = \lambda(\nabla\phi(A))$ for all Borel $A \subseteq \mathbb{R}^n$. Hence

$$\lim_{r \rightarrow 0} \frac{\lambda(\partial\phi(B_r(x)))}{\lambda(B_r(x))} = \lim_{r \rightarrow 0} \frac{\lambda(\nabla\phi(B_r(x)))}{\lambda(B_r(x))} = D_\lambda\mu(x)$$

for all x . By Theorem 3.5.2, $\nabla\phi$ is not differentiable at z .

To complete the proof of Theorem 3.5.1, we need to show that f is a computable function. Theorem 3.4.1 shows that there exists a computable convex function ϕ such that $f = \nabla\phi$. In the following subsections, we prove that under some additional assumptions on μ , ϕ is actually $C^{1,\alpha}$ and thus $\nabla\phi$ is computable.

3.5.1. Computability of $\nabla\phi$. For a given martingale M , we define a computable probability measure μ_M on $[0, 1]^n$ by

$$\mu_M([\sigma]) = \lambda([\sigma]) \cdot M(\sigma) \text{ for all } \sigma \text{ with } |\sigma| = ns \text{ for some } s.$$

Lemma 3.5.3. *Let M be a computable martingale and let μ_M be the corresponding probability measure on $[0, 1]^n$. Let $z \in [0, 1]^n$ and let Z be the binary expansion of z . Suppose, the following two conditions hold:*

- (P1) *the measure μ_M is absolutely continuous, not differentiable at z , and*
- (P2) *$0 < M(\sigma) < C$ for some fixed C and all σ .*

Then the optimal transport map from μ_M to λ is computable.

Proof. We know that there is a computable convex function ϕ such that $\nabla\phi$ is the optimal transfer map of μ onto λ . Define $h(x) = D_\lambda\mu_M(x)$. Then ϕ is an Aleksandrov solution of the following instance of the Monge-Ampère equation:

$$\det D^2\phi = h.$$

Since h is bounded away from both 0 and ∞ , by Theorem 4.13 in [12], ϕ is $C^{1,\alpha}$ for some $\alpha > 0$. Since $\nabla\phi$ is a.e. computable and Hölder continuous, it must be computable. \square

To complete the proof we will describe a construction of a martingale M satisfying the conditions of the previous lemma.

3.5.2. Construction of the martingale M .

Lemma 3.5.4. *Suppose $z \in [0, 1]^n$ is not computably random. There does exist a computable martingale M that satisfies the (P1) and (P2) conditions.*

Proof. Again, let Z denote binary expansion of z . We will modify the construction described in the proof of Theorem 4.2 in [4].

The martingale B constructed in the proof of Theorem 4.2 has the following properties relevant to us:

- (1) $1 \leq B(\sigma) \leq 4$ for all σ . Let's call 1 and 4 the *bounding constants* of B .
- (2) B has two distinct *phases*: the *up phase* (where it increases its capital) and the *down phase* (where it decreases the capital). While B is in the up phase, its capital reaches the value of 3, and while B is in the down phase, its capital reaches the value of 2. The construction guarantees that the capital of B along Z oscillates - that is, B alternates between two phases infinitely often. Let's call 2 and 3 the *oscillation constants* of B .

The martingale B satisfies the property $P2$, but not necessarily the $P1$ property since $D_\lambda \mu_B(z)$ might still exist despite oscillations of B on the binary expansion of z .

The construction can be modified to define a variant of B (let's call it M), that satisfies both properties.

Firstly, note that both bounding and oscillation constants are flexible. In particular, we may assume that oscillation constants of M are some rational numbers $p > q > 1$ and its bounding constants are $q - 1$ and $p + 1$. Also note that when the capital of M reaches the value $> p$ ($< q$), M can maintain its capital at the $\geq p$ ($\leq q$) level for as long as needed.

Before explaining what conditions on M guarantee both $P1$ and $P2$ hold, let us define some notation and recall one geometric fact about basic dyadic cubes.

Let $n \in \mathbb{N}$. Define \mathcal{D}_n to be the set of all (basic) dyadic cubes in \mathbb{R}^n . Let $T_n = \{0, 1/3, 2/3\}^n$. For every $t \in T_n$ define $\mathcal{D}_n^t = \{Q + t : Q \in \mathcal{D}_n\}$ and denote by Z^t the binary expansion of $z + t$.

The following fact is known as the “one third trick”.

Fact 3.5.5. *There is a universal constant $K_n > 0$ such that for any ball $B \subset \mathbb{R}^n$ with radius $r < 1/3$, there is $t \in T_n$ and a cube $Q \in \mathcal{D}_n^t$ containing B whose radius is no more than $K_n \cdot r$.*

(TODO: replace with a variant of Theorem 3.8 from Olli Tappiö's thesis and name the constants appropriately)

A simple consequence of this fact is that there exists $k_n \in \mathbb{N}$, such that for any sufficiently small r the following holds

$$[Z^t \upharpoonright_{sn}] \subseteq B_r(z + t) \subseteq [Z^t \upharpoonright_{sn-k_n}] \quad (3)$$

for some $t \in T_n$ and $s \in \mathbb{N}$.

Thus, for some particular value of t , (3) holds for infinitely many s . Since the operation $f \mapsto f + t$ preserves monotonicity and computability, we may assume (3) holds infinitely often for $t = 0$.

Suppose (3) holds for some s, r and $M(Z \upharpoonright_{sn-k_n}) \leq q$. Let $D_2 = [Z \upharpoonright_{sn}]$, $B = B_r(z)$ and $D_1 = [Z \upharpoonright_{sn-k_n}]$. Note that

$$2^{-k_n} \leq \frac{\lambda(B)}{\lambda(D_1)} \leq 1 \text{ and } 2^{-k_n} \leq \frac{\lambda(D_2)}{\lambda(B)} \leq 1.$$

Then we have

$$\frac{\mu(B)}{\lambda(B)} = \frac{\mu(D_1) - \mu(D_1 \setminus B)}{\lambda(D_1) - \lambda(D_1 \setminus B)} \leq \frac{\mu(D_1)}{\lambda(D_2)} = 2^{k_n} \cdot \frac{\mu(D_1)}{\lambda(D_1)} \leq 2^{k_n} \cdot q.$$

Analogously, assuming $M(Z \upharpoonright_{sn}) \geq p$, we get

$$\frac{\mu(B)}{\lambda(B)} = \frac{\mu(D_2) + \mu(B \setminus D_2)}{\lambda(D_1)} \geq \frac{\mu(D_2)}{\lambda(D_1)} = 2^{-k_n} \cdot \frac{\mu(D_2)}{\lambda(D_2)} \geq 2^{-k_n} \cdot p.$$

Hence the following three conditions imply that $D_\lambda \mu(z)$ does not exist:

- (1) $2^{-k_n} \cdot p - 2^{k_n} \cdot q > 0$,
- (2) $M(Z \upharpoonright_{sn}) \geq p$ and $[Z \upharpoonright_{sn}] \subseteq B_r(z) \subseteq [Z \upharpoonright_{sn-k_n}]$ hold for infinitely many s ,
and
- (3) $M(Z \upharpoonright_{sn-k_n}) \leq q$ and $[Z \upharpoonright_{sn}] \subseteq B_r(z) \subseteq [Z \upharpoonright_{sn-k_n}]$ hold for infinitely many s .

The first condition is trivially met by setting p and q appropriately. To ensure the second condition is met, the martingale M , once it is in the up phase and its capital is $> p$, waits (maintaining the value of its capital $> p$) until it is clear that there are two basic dyadic cubes satisfying 3 (this can be done by checking that the distance between boundaries of two dyadic cubes is large enough).

The third condition can be dealt with in an analogous manner. \square

REFERENCES

- [1] G. Alberti and L. Ambrosio. A geometrical approach to monotone functions in \mathbb{R}^n . *Mathematische Zeitschrift*, 230:259–316, 1999.
- [2] V. I. Bogachev. *Measure theory. Vol. I, II*. Springer-Verlag, Berlin, 2007.
- [3] V. Brattka, J. Miller, and A. Nies. Randomness and differentiability. *Transactions of the AMS*, Forthcoming. <http://arxiv.org/abs/1104.4465>.
- [4] C. Freer, B. Kjos-Hanssen, A. Nies, and F. Stephan. Algorithmic aspects of lipschitz functions. *Computability*, 3(1):45–61, 2014.
- [5] Peter Gács. Uniform test of algorithmic randomness over a general space. *Theoret. Comput. Sci.*, 341(1-3):91–137, 2005.
- [6] M. Hoyrup and C. Rojas. Computability of probability measures and Martin-Löf randomness over metric spaces. *Inform. and Comput.*, 207(7):830–847, 2009.
- [7] R. J. McCann. A convexity principle for interacting gasses. *Adv. Math.* 128, 1:153–179, 1997.
- [8] F. Mignot. Contrôle optimal dans les inéquations variationnelles elliptiques. *J. Funct. Anal.*, (22):130–185, 1976.
- [9] G. Minty. Monotone nonlinear operators on a Hilbert space. *Duke Math J.*, (29):341–346, 1962.
- [10] R. T. Rockafellar and R. J-B Wets. *Variational Analysis*. Grundlehren der Mathematischen Wissenschaften 317. Springer-Verlag, 1997.
- [11] J. Rute. Computable randomness and betting for computable probability spaces. 2012. Submitted.
- [12] C. Villani. *Topics in Optimal Transportation*, volume 58 of *Graduate Studies in Mathematics*. American Mathematical Society, 2003.